GLBL 503: Introduction to Statistics

102 Baker Institute, R 6:00p-9:00p

Dr. Justin Esarey
Assistant Professor of Political Science
Rice University

Office: 108 Herzstein Hall
Office Hours: M T 2:00 – 3:00 PM
E-mail: justin@justinesarey.com
Phone: 678-383-9629

COURSE OBJECTIVES AND LEARNING OUTCOMES

This course introduces students to basic ideas in statistical analysis and the application of all these ideas to substantive questions in political science.

Students will be able to:

1. understand and interpret basic descriptive statistics,
2. understand and interpret linear regression results,
3. understand basic principles of causal inference in experimental data,
4. understand fundamental barriers to causal inference in observational data, and
5. use the R statistical environment to analyze data using basic descriptive statistics and linear regression.

GRADING POLICIES AND ASSIGNMENT DETAILS

Grade Components:

- Homework: 50%
- Final Exam: 50%

Grading Scale:

100%-97%: A+
96.9%-93%: A
92.9%-90%: A-
89.9%-87%: B+
86.9%-83%: B
82.9%-80%: B-
79.9%-77%: C+
76.9%-73%: C
72.9%-70%: C-
69.9%-67%: D+
66.9%-63%: D
62.9%-60%: D-
>59.5%: F

Exams: There will be a final exam in this class. The exam will be a take-home exam. No materials, references, or other students may be consulted during the exam.

Homework: Homework problem sets will be distributed during class. I encourage collaborative work on problem sets: the goal of a homework problem set is to help you learn the material and enable you to perform well on the (non-collaborative!) exam. With that said, simply copying another student's homework answers is not permitted and will be treated as academic dishonesty.

Attendance: Attendance is mandatory in this class. No points are deducted on the basis of an absence alone; however, anything I teach in class is testable material, and not everything I teach in class will be in the course reading material.

COURSE POLICIES

Courtesy: To ensure that everyone has the opportunity to learn without undue distraction, please follow these guidelines:

(1) Turn off your cell phone.
(2) Do not talk during class unless called upon by the instructor.
(3) Do not read newspapers, magazines, or non-class-related internet web sites (e.g., Facebook) during class.
(4) Do not enter class late or leave early unless it is an emergency or you have cleared it with me in advance.

If you violate these rules, I may ask you to leave the classroom. I reserve the right to penalize your grade for repeated inappropriate behavior, up to and including permanent removal from the class.

Late Work: Assignments are due at the date and time I specify for the assignment. Late homeworks will be marked off at 5 percentage points for the first 24 hours late, and an additional 10 percentage points for every subsequent 24 hours late.

Exams may not be taken late. Failing to take the exam at the scheduled time will result in no credit for the exam.

Penalties for late homeworks or missed exams may be waived under the following three circumstances:

(1) a death in the immediate family (parent, spouse, sibling, or child) within two weeks before the assignment due date;
(2) in the event of an unforeseeable medical emergency affecting you, your spouse, or your child; or
(3) participation in a Rice-sponsored academic or sporting event.

In the event of (1) or (3), you must give me **at least 24 hours advance notice and preferably more** (via e-mail or a phone call) that you will miss the exam/homework, or it may not be made up. I may require supporting documentation. All penalty waivers are at the discretion of the instructor. Under these circumstances, I will extend your homework due date or schedule you a make-up exam time. **Important note: conflicts with a work schedule, non-academic trip, or vacation are not a valid reason to miss an exam or any other assignment** and cannot be the basis for a penalty waiver.

Honor Code/Academic Misconduct: All forms of academic misconduct will be handled according to the Rice University Honor Code. Details on the Honor Code are available at http://honor.rice.edu/honor-system-handbook/.

If you ever have any questions about what you should do to stay within the honor code on a particular assignment, **contact me with your question and I can assist you**. I cannot guarantee a timely response unless you contact me at least 24 hours in advance of the time the assignment is due.

Students with Disabilities: If you have a disability and require accommodation in this class, please contact me as soon as possible (within the first two weeks of class) to discuss these accommodations. You will also need to contact the Disability Support Services Office (telephone extension: 5841) in the Allen Center.

Syllabus Change Policy: All policies of this syllabus may be changed by Prof. Esarey with advance notice.

COURSE MATERIALS

Required Texts:

- Angrist and Pischke. 2015. *Mastering 'Metrics: The Path from Cause to Effect.* Princeton University Press.
- Dalgaard. 2008. *Introductory Statistics with R.* Springer.
- Wheelan. 2013. *Naked Statistics: Stripping the Dread from the Data.* W.W. Norton.

Other readings are available on the OWL-Space website.

Software: This course will teach material primarily through R. R is free and available for Windows and Macintosh from http://cran.r-project.org/.

I recommend installing The RStudio IDE environment for writing and executing code in R; RStudio is also free and available for Windows and Macintosh at http://www.rstudio.com.

All students must have a valid Rice e-mail address and login (and access to the OWL-space website) to participate in this course.

COURSE OUTLINE AND ASSIGNED READINGS

1) Introduction to R and basic descriptive statistics

Readings:

a. Wheelan, Chapters 1, 2, 3
b. Jo Hardin, "An R Tutorial." URL: http://goo.gl/UzzSL7
c. Dalgaard, Chapters 1, 2

2) Probability and more descriptive statistics

Readings:

a. Wheelan, Chapters 5, 6, and 8
b. Dalgaard, Chapter 3-4
c. Angrist and Pischke, Chapter 1 Appendix ("Mastering Inference")
d. **Optional:** Wooldridge, *Introductory Econometrics*, Appendices B and C

3) The logic of experimentation and causal inference

Readings:

a. Angrist and Pischke: Chapter 1
b. Wheelan, Chapter 9
c. Dalgaard, Chapter 5

4) Regression I

Readings:

a. Wheelan, Chapter 11
b. Angrist and Pischke: Chapter 2
c. Dalgaard, Chapter 6

5) Regression II

Readings:

a. Wheelan, Chapters 12 and 13
b. Maindonald and Braun, "Multiple Linear Regression." Chapter 6 in *Data Analysis and Graphics Using R: An Example-Based Approach.* Cambridge University Press. Available on OWL-Space.
c. Alan Abramowitz, "Forecasting the 2008 Presidential Election with the Time-for-Change Model." *PS: Political Science and Politics* 41: 691-695. URL: http://goo.gl/DLckXv.

6) Instrumental Variables Designs

Readings:

a. Angrist and Pischke: Chapter 3
b. Hyland, "A Simple Instrumental Variables Problem." URL: http://goo.gl/3psv9e.
c. Acemoglu, Johnson, and Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *The American Economic Review* 91(5): 1369-1401. URL: http://goo.gl/57CUYJ.

7) Difference-in-Difference Designs

Readings:

a. Angrist and Pischke: Chapter 5
b. Holbein and Hillygus. 2015. "Making Young Voters: The Impact of Preregistration on Youth Turnout." *American Journal of Political Science* forthcoming. URL: http://goo.gl/oWTkFH.

8) Crosstabs and the χ^2 test

Homework #4 distributed.

Readings:

a. Kellstedt and Whitten, Ch. 7, Section 7.4.1
b. Maindonald and Braun, Chapter 4, Section 4.3.
c. Second look: Alan S. Gerber and Donald P. Green, "Do Phone Calls Increase Voter Turnout? A Field Experiment." *Public Opinion Quarterly* 65 (Spring 2001), 75-85; see especially p. 79 for data for a chi-squared test. URL: http://goo.gl/KAdNMW.

9) Correlation

Readings:

a. Kellstedt and Whitten, Ch. 7, Section 7.4.3
b. Amy Caiazza, "Does Women's Representation in Elected Office Lead to Women-Friendly Policy?" *Women and Politics* 26 (2004), 35-70. URL: http://goo.gl/C417VH.

********** **EXAM TWO CUTOFF** **********

10) Regression I: theory and basics

Readings:

a. Kellstedt and Whitten, Ch. 8., Sections 8.1-8.4.
b. Maindonald and Braun, Ch. 5, Section 5.1 and 5.4.

11) Regression II: inference, multiple regression, and interpretation

Homework #5 distributed.

Readings:

a. Kellstedt and Whitten, Ch. 9
b. Maindonald and Braun, Ch. 5, Section 5.3 and Ch. 6, Section 6.1-6.3.

12) Regression III: applied examples and categorical explanatory variables

Readings:

a. Kellstedt and Whitten, Ch. 10, Sections 10.1-10.2.
b. Maindonald and Braun, Ch. 7, Sections 7.1-7.3
c. John J. Donohue III and Steven D. Levitt, "The Impact of Legalized Abortion on Crime," *The Quarterly Journal of Economics* 46 (May 2001): 379-420. URL: http://goo.gl/gtLIlf.
d. Alan Abramowitz, "Forecasting the 2008 Presidential Election with the Time-for-Change Model." *PS: Political Science and Politics* 41 (October 2008): 691-695. URL: http://goo.gl/DLckXv.

13) Regression IV: assumptions underlying regression

Homework #6 distributed.

Readings:

a. Kellstedt and Whitten, Ch. 8, Section 8.5, and Ch. 10, Section 10.4-10.5
b. Maindonald and Braun, Ch. 5, section 5.2; Ch. 6, Sections 6.4-6.6; Ch. 7, Section 7.4-7.5.

14) LPM, logit and probit I: studying binary dependent variables

Readings:

a. Kellstedt and Whitten, Ch. 11, Sections 11.1-11.2
b. Maindonald and Braun, Ch. 8, Sections 8.1-8.2.

15) Logit and probit II: plotting and interpreting marginal effects

Homework #7 distributed.

Readings:

a. G. King, M. Tomz, and J. Wittenberg, "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science* 44 (2000): 347-361. URL: http://goo.gl/9StFf8.
b. D. Sunshine Hillygus and Todd G. Shields, "Moral Issues and Voter Decision Making in the 2004 Presidential Election." *PS: Political Science and Politics* (April 2005), 201-209

16) Interaction effects

Readings:

a. Kellstedt and Whitten, Ch. 10, Section 10.3
b. T. Brambor, W. R. Clark, and M. Golder, "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (2006): 63-82. URL: http://goo.gl/iQrY5a.
c. Justin Esarey and Gina Chirillo, "'Fairer Sex' or Purity Myth? Corruption, Gender, and Institutional Context." *Politics and Gender* (forthcoming). URL: http://goo.gl/2yusN0.